Visual dominance and multisensory integration changes with age

Diaconescu, Andreea Oliviana ; Hasher, Lynn ; McIntosh, Anthony Randal

Abstract: Objects comprise of visual and auditory signatures that arrive through distinct sensory channels. Exposure to cross-modal events sets up expectations about what a given object most likely "sounds" like, and vice versa, thereby facilitating detection and recognition. Whereas episodic and working memory functions decline with age, the extent to which multisensory integration processes change with age remains an open question. In the present study, we examined whether multisensory integration processes play a compensatory role in normal aging. Magnetoencephalography recordings of semantically-related cross-modal and unimodal auditory and visual stimuli captured the spatiotemporal dynamics of multisensory responses in young and older adults. Whereas sensory-specific regions showed increased activity in response to cross-modal compared to unimodal stimuli 100 ms after stimulus onset in both age groups, posterior parietal and medial prefrontal regions responded preferentially to cross-modal stimuli between 150 and 300 ms in the older group only. Additionally, faster detection of cross-modal stimuli correlated with increased activity in inferior parietal and medial prefrontal regions 100 ms after stimulus onset in older compared to younger adults. Age-related differences in visual dominance were also observed with older adults exhibiting significantly larger multisensory facilitation effects relative to the auditory modality. Using structural equation modeling, we showed that age-related increases in parietal and medial prefrontal source activity predicted faster detection of cross-modal stimuli. Furthermore, the relationship between performance and source activity was mediated by age-related reductions in gray matter volume in those regions. Thus, we propose that multisensory integration processes change with age such that posterior parietal and medial prefrontal activity underlies the integrated response in older adults.

DOI: https://doi.org/10.1016/j.neuroimage.2012.09.057

Visual dominance and multisensory integration changes with age

Andreea Oliviana Diaconescu[3]
Lynn Hasher[1,2]
Anthony Randal McIntosh[1,2]

1. Rotman Research Institute, Baycrest Centre, Toronto, Ontario, Canada M6A 2E1
2. Department of Psychology, University of Toronto, Ontario, Canada M8V 2S4
3. Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zürich/ETH Zürich, Switzerland

Corresponding Author: Andreea Oliviana Diaconescu
Translational Neuromodeling Unit (TNU)
Institute for Biomedical Engineering
University of Zürich/ETH Zürich, Switzerland

Email: andreeaoliviana.diaconescu@uzh.ch

Manuscript Information:

Keywords: aging, multisensory facilitation, auditory, visual, MEG, event-related SAM analysis

Disclosure: we report no conflicts of interest

Acknowledgments: This research was supported by grants from the Canadian Institutes of Health Research and the JS McDonnell Foundation to A.R.M. and the Canadian Institutes of Health Research to L.H. A.O.D. was supported by the Natural Sciences and Engineering Research Council of Canada.

Abstract

Objects comprise of visual and auditory signatures that arrive through distinct sensory channels. Exposure to cross-modal events sets up expectations about what a given object most likely "sounds" like, and vice versa, thereby facilitating detection and recognition. Whereas episodic and working memory functions decline with age, the extent to which multisensory integration processes change with age remains an open question. In the present study, we examined whether multisensory integration processes play a compensatory role in normal aging. Magnetoencephalography recordings of semantically-related cross-modal and unimodal auditory and visual stimuli captured the spatiotemporal dynamics of multisensory responses in young and older adults. Whereas sensory-specific regions showed increased activity in response to cross-modal compared to unimodal stimuli 100ms after stimulus onset in both age groups, posterior parietal and medial prefrontal regions responded preferentially to cross-modal stimuli between 150 and 300ms in the older group only. Additionally, faster detection of cross-modal stimuli correlated with increased activity in inferior parietal and medial prefrontal regions 100ms after stimulus onset in older compared to younger adults. Age-related differences in visual dominance were also observed with older adults exhibiting significantly larger multisensory facilitation effects relative to the auditory modality. Using structural equation modeling, we showed that age-related increases in parietal and medial prefrontal source activity predicted faster detection of cross-modal stimuli. Furthermore, the relationship between performance and source activity was mediated by age-related reductions in gray matter volume in those regions. Thus, we propose that multisensory integration processes change with age such that posterior parietal and medial prefrontal activity underlies the integrated response in older adults.

1. Introduction

Changes in basic auditory and visual processing occur with age. Besides age-related reductions in visual acuity (Cerella, 1985; Spear, 1993), peripheral auditory processes, including the efficiency of temporal and spectral resolution (cf., Alain, Dyson, & Snyder, 2006; Alain, Ogawa, & Woods, 1996; Pichora-Fuller, Schneider, & Daneman, 1995; Schneider, Speranza, & Pichora-Fuller, 1998) also decline with age. In addition to perceptual changes, a recent qualitative examination of age-related changes in cognition suggested that while semantic knowledge remains stable with age (Craik & Bialystok, 2006), fluid intelligence, which is partly determined by attentional regulation, decreases in efficiency, speed, and complexity (Salthouse, 2007).

Attentional regulation is influenced by many factors including reductions in the ability to suppress distracting information in a task-relevant context (Hasher, Stoltzfus, Zacks, & Rypma, 1991; Hasher, Zacks, & May, 1999). Older adults are more vulnerable than young adults to the disruptive effects of concurrent distraction as can be seen in basic perceptual speed (Lustig, Hasher, & Tonev, 2006), visual search (Scialfa, Esau, & Joffe, 1998), selective attention (Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993), controlled memory search (Hartman & Hasher, 1991), sustained attention (Bunce, Warr, & Cochrane, 1993), verbal problem solving (May et al., 1999), Stroop (Spieler, Balota, & Faust, 1996; West & Alain, 1999, 2000), and flanker tasks (Zeef, Sonke, Kok, Buiten, & Kenemans, 1996). Presumably because they are not sufficiently dampened, task- irrelevant stimuli may receive a richer representation and become more strongly linked to target stimuli.

Potentially related to reduced inhibitory control, responses to temporally coincident cross-modal stimuli may be less specific in older adults, thus leading to enhanced integration across sensory modalities. A recent study showed that a temporally synchronized tone facilitates the detection of a visual target, thereby reducing the effects of concurrent distraction in older compared to younger adults (Campbell, Al-Aidroos, Fatt, Pratt, & Hasher, 2010). Older adults benefited more from multisensory presentations than younger adults as evidenced by increased saccadic trajectory deviations away from task-irrelevant distractors during cross-modal compared to unimodal presentations. Furthermore, previous studies using concurrent presentations of auditory and visual stimuli in simple detection tasks showed enhanced cross-modal facilitation in older compared to younger adults (Diederich, Colonius, & Schomburg, 2008; Laurienti, Burdette, Maldjian, & Wallace, 2006; Peiffer, Mozolic, Hugenschmidt, &

Laurienti, 2007). On the other hand, although older adults benefit more than young adults from temporally- and semantically-congruent cross-modal information, they also show greater impairment in integrating incongruent cross-modal inputs (Poliakoff, Coward, Lowe, & O'Boyle, 2007; Poliakoff, Shore, Lowe, & Spence, 2006; Setti et al., 2011).

Multisensory integration is an important process to examine in aging because it may be one of the few contexts in which reduced inhibitory control and enhanced binding across temporally coincident stimuli can confer an additional benefit to older adults. Visual sensory dominance effects may also be stronger during late adulthood. Recently, Guerreiro et al. (2010) proposed that older adults are more sensitive to interference from task-irrelevant visual stimuli than auditory ones. The group performed a large meta-analysis of previous studies that examined inhibitory control in aging, and concluded that, in older adults, task-irrelevant auditory suppression during visual attention was more successful than task-irrelevant visual suppression during auditory attention tasks.

While there is growing evidence that older adults respond faster to cross-modal stimuli (Campbell et al., 2010) and that integrating congruent information across sensory channels may differentially improve response latencies in aging (Laurienti et al., 2006; Peiffer et al., 2007), the neural dynamics that can explain age differences in multisensory responses are still unknown.

To examine the spatiotemporal dynamics underlying multisensory responses and age differences in sensory dominance, we collected magnetoencephalography (MEG) recordings during the presentation of complex sounds and semantically-related black-and-white line drawings of animate and inanimate objects in both young and older adults. We used visual images presented along with congruent complex sounds to create multisensory associations that were previously established in the course of an individual's learning history. As cross-modal congruence facilitates object detection and recognition because individuals exhibit precise expectations about what a given object most likely "sounds" like, for example, we predicted that multisensory facilitation effects would be larger using complex objects instead of simple stimuli. Furthermore, we also predicted that sensory modality dominance effects would be more pronounced for complex as compared to simple audiovisual objects given recent evidence of visual dominance in behavioural studies using complex audiovisual objects. For example, Chen and Spence (2010) recently showed that semantically congruent complex sounds that were presented concurrently with visual targets improved while semantically-incongruent sounds impaired participants' ability to identify visual targets.

In the present context, the multisensory neural response represents the process whereby unisensory signals are combined to form a unique signal that is specifically associated with the cross-modal stimulus and is significantly distinct from the sum of the responses evoked by the modality-specific component stimuli (cf. Stein et al. 2010). Previous event-related potential (ERP) studies in young adults showed that sensory-specific regions, namely primary auditory, primary visual, ventral occipito-temporal and superior temporal cortices displayed increased activity in response to temporally- and semantically-congruent cross-modal inputs (Cappe, Thut, Romei & Murray, 2010; Raij et al., 2010; Teder-Sälejärvi, Russo, McDonald & Hillyard, 2005). Furthermore, intracranial electrophysiological studies (Molholm et al., 2006; Moran, Molholm, Reilly, & Foxe, 2008), magnetoencephalography (MEG) studies (Diaconescu, Alain, & McIntosh, 2011; Raij et al., 2010) and functional neuroimaging studies in humans (Baumann & Greenlee, 2007; Bishop & Miller, 2008; Calvert, Hansen, Iversen, & Brammer, 2001; Grefkes, Weiss, Zilles, & Fink, 2002; Macaluso, George, Dolan, Spence, & Driver, 2004) showed that cross-modal stimuli did not only elicit increased activity in sensory-specific cortices, but also activated a distinct network of posterior parietal brain regions, including the inferior parietal sulcus (IPS), the inferior parietal lobule (IPL), and the superior parietal lobule (SPL).

We predicted that older adults will show enhanced integration between temporally- and semantically-congruent auditory and visual stimuli. As a result, they may exhibit faster response times (RTs) to cross-modal compared to unimodal stimuli as well as larger multisensory responses in sensory-specific and multimodal brain regions. On the other hand, as a result of enhanced visual dominance effects, older adults may show RT facilitation to cross-modal stimuli relative to unimodal auditory ones, but the presence of an auditory target may not speed up the detection of the visual target.

Furthermore, we also accommodate the possibility that multisensory responses are the same in both young and old, but that the dynamics or the reliance on a particular brain area may differ between young and older participants. We propose that whereas young adults selectively recruit sensory-specific and posterior parietal regions to integrate cross-modal, semantically-congruent stimuli (Diaconescu et al., 2011), older adults may engage additional prefrontal regions while performing the same tasks. This hypothesis is supported by the theory of cortical dedifferentiation, which posits that healthy aging is accompanied by decreased neuronal specificity in prefrontal cortices (for a review, see Park & Reuter-Lorenz, 2009). As a result of

overall reduction in gray matter volume with age (cf., Gunning-Dixon & Raz, 2000; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003), prefrontal cortical regions that were functionally specialized, with age, respond more similarly across a variety of cognitive states. For example, Cabeza, Anderson, Locantore and McIntosh (2002) proposed that bilateral recruitment of medial prefrontal cortices is an adaptive, compensatory neural mechanism in older adults because activity in these regions predicted superior performance on a source memory task in older compared to younger participants.

2. Materials and Methods

2.1 Participants

Fifteen young adults (7 males, 20-29 years, mean age ± s.d., 23.46 ± 3.09) and sixteen older adults (8 males, 66-78 years, 69.93 ± 4.76) in each group with an average of 16.5 years of education and a mini mental status exam score of 28 and higher (Folstein, Folstein, & McHugh, 1975) participated in the present study. All participants were right-handed with healthy neurological histories and normal to corrected-to-normal vision. All volunteers were audiometrically screened to determine hearing thresholds for each ear separately using both low and high frequency tones (800 Hz and 1200 Hz) to account for common sensory deficits in aging, namely presbycusis. Volunteers whose hearing thresholds exceeded 15dB hearing level (HL) were excluded from participation as that was considered to be below normal levels. The young adults who participated in the study had average hearing thresholds of 2dB (range 0-8dB), and older adults had average hearing thresholds of 10dB (range 5-15dB). Hearing thresholds significantly differed between young and older adults [$t(1, 29) = 11.13$, $p < 0.001$].

To assess cognitive functions, three neuropsychological tests were administered during in-person interviews. Concentration and attention were measured using the Short Blessed Test (Katzman et al., 1983), vocabulary using the Shipley Institute for Living Scale Test (Shipley, 1991), and memory and motor function using Folstein's Mini-Mental Status examination (Folstein et al., 1975). All participants had a mini mental status exam score of 28 and higher (Folstein et al., 1975) and met the cut-off of 6 on the SBT (Katzman et al., 1983), suggesting normal cognitive functions. Older adults, however, scored significantly higher than younger adults on vocabulary testing [$F(2, 58) = 13.15$, $p < 0.05$] supporting previous findings of improved crystallized knowledge and vocabulary with age (Grady & Craik, 2000; Park et al., 2002; Salthouse, 2007).

The joint Baycrest Centre-University of Toronto Research Ethics Committee approved the study and the rights and privacy of the participants were maintained. All participants gave formal informed consent before the experiment and received monetary compensation.

2.2 Stimuli

Stimuli were selected to have semantically congruent auditory and visual representations. Two types of stimuli, animate and inanimate, were used in the study. Items were selected from 4 distinct categories: (1) animals, (2) musical instruments, (3) automobiles, and (4) household objects. The first category of stimuli was labeled as "animate", while the remaining 3 categories were considered "inanimate" objects.

Black-and-white line drawings, selected from the Snodgrass and Vanderwart (1980) database of visual stimuli, served as the visual stimuli. All visual stimuli were matched according to size (in pixels), brightness, and contrast. Semantically-related non-speech, complex sounds were matched in terms of loudness by taking the average of the root mean square values across all complex sounds. Each sound was assigned the mean amplitude; thus, louder sounds were reduced, while softer ones were amplified. The sounds were delivered binaurally at an intensity level of 60dB HL based on the audiometric mean across both ears.

Visual and auditory unimodal stimuli were then paired to create cross-modal stimuli that were matched semantically (e.g., picture of a lion paired with the sound of a roar or picture of an ambulance car paired with a siren). Thus, three stimulus types were employed: (1) auditory unimodal (A), (2) visual unimodal (V), and (3) cross-modal semantically congruent (AV+) and semantically incongruent (AV-).

To ensure that the complex sounds were easily nameable and identifiable, we assessed accuracy values and RTs for each stimulus exemplar in an initial behavioural pilot. Five young and older adults (mean ages 26 and 64, respectively), participated in the pilot study. Complex sounds were excluded if detection accuracy levels fell below 75% and RTs exceeded two standard deviations above the mean RT values for each individual subject. Following behavioural testing, we asked participants to rate complex sounds based on their recognizability and classifiability. On the basis of the behavioural findings and the post-experiment questionnaire results, we excluded several complex sounds along with their visual counterparts. Thus, for each animate or inanimate category, 30 different exemplars from each sensory modality (auditory and visual) were selected because they could be unambiguously categorized (Supplementary Table 1). In total, 60 animate stimuli (30 auditory and 30 visual) and 60 inanimate stimuli (30 auditory and 30 visual) were used.

2.3 Procedure

Each stimulus or stimulus pair was presented for 400ms; for the auditory stimuli, the 400ms interval also included a 5ms fall and rise time. The time interval between the end of the stimulus presentation and the beginning of the next trial was either 2, 3 or 4s (equiprobable). See Figure 1 for an illustration of the paradigm.

Two tasks were used to assess multisensory responses, namely simple detection and semantic classification. In the simple detection task, participants were instructed to respond (left index finger response) as quickly as possible to any stimulus type: unimodal A, V and AV+ or AV-. In the semantic classification task, participants made animacy or inanimacy judgments for all stimulus presentations: unimodal A, V and AV+ (left index and middle finger responses for animate and inanimate judgments, respectively). Please note that only congruent cross-modal presentations were included in the semantic classification task, such that the sound of a lion's roar was paired with an image of a lion, for example. Forty stimuli from each trial type (i.e., A, V, AV+ and AV-) were presented in each block for 400ms in a total of 160 trials in the simple detection task and 120 trials in the semantic classification task. Presentation software (version 10.3; Neurobehavioural Systems, Inc.; http://www.neurobs.com/) was used to control visual and auditory stimulus delivery and to record participants' response latency and accuracy.

As the behavioural and neural responses did not differ between congruent and incongruent presentations (i.e., AV+ and AV- trial types) in the simple detection task (cf. Table 1a and b), we compared unimodal trial types to AV+ trial types only. This procedure also ensured that we used an equal number of cross-modal presentations in both simple detection and semantic classification tasks.

2.4 Structural MRI Acquisition and Analysis

Source maps were computed on each participant's structural magnetic resonance imaging (MRI) scan. The scans were acquired using a 3.0T Siemens Tim MAGNETOM Trio MRI scanner (Software level Syngo MR, Siemens Medical, Germany) with 12-channel head coil.

Estimates of cortical thickness and gray or white matter volumes were also obtained using FreeSurfer (Dale, Fischl, & Sereno, 1999). For each subject, gray matter, white matter, and all other non-cortical structures were segmented and a triangular mesh was used to measure

the distance from the pial surface to the gray matter/white matter boundary for each hemisphere. Furthermore, a quality control inspection assessed for gross structural abnormalities, accuracy of registration, and presence of artifacts. Volumes of white matter and gray matter and mean cortical thickness were computed for each lobe and for both the left and the right hemispheres.

2.5 MEG recordings

The magnetoencephalogram was recorded in a magnetically shielded room at the Rotman Research Institute, Baycrest Centre, using a 151-channel whole head neuro-magnetometer (OMEGA, VSM Medtech Inc., Vancouver, Canada). Participants sat upright, and viewed the visual stimuli on a back projection screen that subtended approximately 30 degrees of visual angle when seated 70cm from the screen. With respect to the visual presentations, the MEG collection was synchronized to the onset of each stimulus by recording the luminance change of the screen with a photodiode. Binaural auditory stimuli were presented at 60dB HL via OB 822 Clinical Audiometer through ER30 transducers (Etymotic Research, Elk Grove, USA) and connected with 1.5 m matched plastic tubing and foam earplugs to the participant's ears. With respect to the auditory stimuli, the MEG data collection was synchronized to the onset of the auditory sound envelope.

Participants' head positions within the MEG were determined at the start and end of each recording block using indicator coils placed on nasion and bilateral pre-auricular points. These fiducial points established a head-based Cartesian coordinate system for pre-processing and analysis of the MEG data.

2.6 MEG Pre-processing

Neuromagnetic activity was sampled at a rate of 1250 Hz, and was recorded continuously in four experimental blocks (i.e., with each task performed twice) of 15 minutes recording time each. Third gradient noise correction was applied to the continuous MEG data. Afterwards, the MEG data were parsed into epochs including a 200ms pre- and 1000ms of post-stimulus activity window, and D.C. offsets were removed from the entire epoch. Finally, MEG data were band-pass filtered between 0.1 and 55 Hz and averaged across all trial types. A principal component analysis (PCA) was performed on each epoch and components larger than 2.0 pT at any time point were subtracted from the data (Okada, Jung, & Kobayashi, 2007). This preprocessing step effectively removed large artifacts caused by eye-blinks. All participants'

structural MRIs and MEG source data were spatially normalized to the Talairach standard brain using Analysis of Functional Neuroimaging software (AFNI; Cox, 1996).

2.7 MEG data analysis

2.7.1 Event-related SAM analysis

We used the synthetic aperture magnetometry (SAM) technique to analyze the source of the measured magnetic field (Robinson & Vrba, 1998; Sekihara, Nagarajan, Poeppel, Marantz, & Miyashita, 2001). SAM minimizes power or the variance of the measured MEG signals such that signals emitted from sources outside each specified voxel are suppressed (Brookes et al., 2007; Cheyne, Bostan, Gaetz, & Pang, 2007). This enables one to display simultaneously active sources at multiple sites, provided that they are not perfectly synchronized.

To obtain spatial precision without integrating power over long temporal windows, we used an event-related version of the SAM analysis technique introduced by Cheyne, Bakhtazad, and Gaetz (2006) to identify evoked brain responses from un-averaged, single trial data. Similar to previous beamforming approaches, the event-related SAM analysis uses the individual trials of each condition and the forward solution for modeling optimal current direction to calculate a spatial filter for each voxel using the minimum-variance beamforming algorithm (Cheyne, Bakhtazad, & Gaetz, 2006).

The spatial filter included 72 brain regions of interest adapted from the regional map coarse parcellation scheme of the cerebral cortex proposed by Kotter and colleagues (Bezgin, Wanke, Krumnack, & Kotter, 2008; Kotter & Wanke, 2005). Please refer to Supplementary Table 2 for a complete listing of the brain regions used with their respective Talairach coordinates and anatomical labels. Each brain region was defined by a three-dimensional position vector and consisted of a unique set of sensor coefficients that constituted a weighting matrix. The MEG data was then projected through this spatial filter to give a measure of current density, as a function of time, in the target brain region. As this source time series was calculated using a weighted sum of the MEG sensors, it had the same millisecond time resolution as the original MEG sensor data.

Source activity at each time point is normalized by the noise power, resulting in a quantity described in units of pseudo-z scores (Robinson & Vrba, 1998). Noise power is essentially uncorrelated random noise that is identical across all channels. This modification is equivalent to normalizing the weights by the noise variance. Thus, the sub- or super-additive

nature of multisensory interactions may not be adequately captured with this technique because the source localization is biased in the presence of both weak and dominant sources. If a combination of weak and dominant sources is present, the beamformer may localize weak sources less accurately (Quraan & Cheyne, 2010). Therefore, weaker sources will have reduced amplitudes compared to more dominant ones making it difficult to draw conclusions about sub- or supra-additive nature of multisensory interactions.

To enhance the spatial precision of this technique, the participants' structural MRIs were used to constrain the event-related SAM images to each participant's individual MRI and to allow for spatial normalization and group averaging in stereotaxic space. The individual functional maps were overlaid on each participant's MRI based on co-registration with the indicator coils initially placed on the nasion and bilateral pre-auricular points. The functional data were then transformed to the standard Talairach-Tournoux space using the same transformation applied to the structural MRI using AFNI (Cox, 1996).

2.7.2 PLS Analysis

We used partial least squares (PLS; Krishnan, Williams, McIntosh, & Abdi, 2011; Lobaugh, West, & McIntosh, 2001) to examine neuromagnetic brain activity across all 72 brain regions of interest as a function of age group and task. The term "partial least squares" refers to the computation of an optimal squares fit to *part* of a covariance structure that is attributable to the experimental manipulations or that relates to a given outcome measure. PLS applied to MEG data is conceptually analogous to the analysis of MEG difference waveforms, because it identifies task-related differences in amplitude across all MEG sources by deriving the optimal least squares contrasts that code for the task differences. Because PLS performs this computation across the entire dataset in time and space simultaneously, there is no need to specify *a priori* MEG sources or time intervals for the analysis.

In the mean-centering approach, trials within each experimental condition were averaged and then expressed as a source-by-source deviation from the grand mean across the entire experiment. Singular value decomposition (SVD) was then applied to the mean-centered deviation matrix.

Mathematically, SVD re-expresses this matrix as a set of orthogonal singular vectors or latent variables (LVs), the number of which is equivalent to the total number of conditions. Each LV contains a pair of vectors relating brain activity to the experimental design. For each LV, the two vectors are linked by a singular value, which is the covariance between the two

blocks of data, namely brain activity and experimental design, and indicates the proportion of cross-block covariance that is accounted for by each LV. The two vectors mentioned above reflect a symmetrical relationship between the components of the experimental design most related to the *differing signals* in the MEG sources on one hand (i.e., task saliences), and the spatiotemporal pattern of MEG sources related the identified experimental design components on the other (i.e., source saliences).

For part of the analyses discussed in this study, we used a non-rotated version of mean-centered PLS, in which *a priori* contrasts restricted the spatiotemporal patterns derived from PLS. This version of PLS has the advantage of allowing direct assessment of hypothesized experimental effects. To assess multisensory responses in the simple detection condition, we contrasted cross-modal congruent trial types to both unimodal auditory and unimodal visual trials trial types. To capture both multisensory responses and age differences we used the following contrasts: (i) crossmodal versus unimodal auditory and unimodal visual trial types in both age groups: [1 -1 -1; 1 -1 -1], and (ii) the interaction between trial types and group assignment: [-1 1 1; 1 -1 -1]. Task differences were examined across the entire epoch, including the pre-stimulus baseline (-0.2 sec) and the post-stimulus interval (1 sec).

2.7.3 Behaviour PLS Analysis

Behaviour PLS was used to examine the correlations between RTs and neuromagnetic activity across all 72 brain regions. The configuration of the data matrix was the same as for the mean-centered approach. However, in contrast to mean-centering the matrix, the correlation between the behaviour measures and the MEG signal at each source was computed across subjects and within each trial-type. The resulting matrix represented a within-task brain-behaviour correlation matrix. SVD applied to this brain-behaviour correlation matrix produced three output matrices. Similar to the mean-centered approach, the three output matrices included the (1) source saliences, (2) singular values, and (3) task saliences. The variations across task saliences, however, indicated in this case whether a given LV represented a similarity or a difference in brain-behaviour correlations across conditions. The source saliences reflected the corresponding brain-behaviour correlation pattern across space (expressed across a collection of MEG sources) and time (expressed across all time points included in the analysis).

2.7.4 Statistical Assessment

Two complementary re-sampling techniques were employed. First, permutation tests assessed whether the task saliences represented by the given LV were significantly different

from random noise. This was accomplished using sampling without replacement, and reassigning the order of the conditions to each subject. Second, the reliability of each source contribution to the LV was assessed using a bootstrap estimation of standard errors for the MEG source saliences. The primary purpose of the bootstrap estimation was to determine the time points of the source waveforms that showed reliable experimental effects across subjects. The use of bootstrap estimation of standard errors eliminates the need to correct for multiple comparisons because the source saliences were calculated in a single mathematical step, on the whole brain at once (McIntosh et al., 1996; McIntosh & Lobaugh, 2004). Statistical evaluation of task effects was performed using an optimal number of 500 permutations (cf., Nichols & Holmes, 2002) and 300 bootstrap iterations (cf., Efron & Tibshirani, 1986; McIntosh, Bookstein, Haxby, & Grady, 1996).

2.8 Structural Equation Modeling

In order to determine the neural mechanisms that underlie the age-differences in multisensory responses, we assessed the directional relationships between cross-modal RT facilitation and the structural and functional measures (i.e., task-related MEG source activity) using structural equation modeling (SEM) in the two age groups. SEM analysis was computed across the two age groups and in both auditory and visual modalities. The models were compared statistically to test for condition-specific differences in path coefficients.

Using SEM, the relationship between three sets of variables was examined. Firstly, MEG source activity was quantified by extracting subject-specific brain scores from the simple detection and semantic classification conditions in auditory and visual modalities, respectively. Brain scores refer to the degree to which each participant expresses the task-dependent contrast identified by a given LV from the hypothesis-driven, non-rotated PLS results (in this case, the interaction between group and task; please refer to the Results section for the details of how this measure was obtained). Brain scores were computed by multiplying the source saliences by the original data. Secondly, the multisensory facilitation index ($RT_{unimodal} - RT_{cross\text{-}modal}$) in auditory and visual modalities was averaged across both simple detection and semantic classification conditions. Finally, the third variable included a composite measure of the gray and white matter volumes in young and older adults. This composite measure was obtained by extracting subject-specific brain scores from the mean-centered PLS analysis, which examined age-differences in gray and white matter volumes across cortical regions in temporal, occipital, parietal and frontal lobes and sub-cortical brain regions, including the thalamus, the striatum, the

cingulate cortex and the hippocampus (please refer to the Results section for the details of how this measure was obtained).

SEM analysis was performed with Amos 19.0 (Amos Development Corp., Meadville, USA) using a maximum likelihood estimation. Statistical inferences about group differences were based on a hierarchical model approach. This approach compared an alternative model, in which all connections were allowed to vary between the two groups and the two conditions, to a null model in which all path coefficients were constrained to be the same across the two age groups and the two conditions, detection and classification.

The χ^2 goodness-of-fit statistic was used to assess the model's ability to reproduce the original correlation matrix. The difference in the fit between null and alternative models (χ^2_{DIFF}) was examined with the degrees of freedom equal to the difference between the degrees of freedom in the constrained and the free models. The χ^2_{DIFF} test is a hierarchical test that examines whether a modification to the model leads to significant improvements in the goodness of fit of the model (McIntosh & Gonzalez-Lima, 1994; Protzner & McIntosh, 2006). If the χ^2 value for the null model is significantly larger than that of the alternative model, then the path coefficients that varied between conditions and groups are statistically distinct (Protzner & McIntosh, 2006). The stability of the path coefficients across subjects in each group was assessed using bootstrap resampling.

3. Results

3.1 Behavioural Performance: Accuracy and RT Trends

Mean and standard deviation values across the two dependent variables, RTs and performance accuracy values, are included in Tables 1a and b for young and older adults in the simple detection and the semantic classification conditions. In the simple detection condition, performance accuracy was above 90% in both groups. However, in the semantic classification condition, accuracy was significantly lower in unimodal auditory trials compared to bimodal ones [$t(1,29) = 11.69$, $p < 0.0001$, $r^2 = 0.825$] and unimodal visual ones [$t(1,29) = 12.63$, $p < 0.0001$, $r^2 = 0.846$].

As an additional factor in the analysis, we also grouped stimuli types into animate and inanimate categories, and examined whether there was a significant difference in performance accuracy between the two stimulus categories. Participants in both groups made significantly more errors while classifying complex sounds, and in both groups, no differences in percent errors were observed between animate and inanimate trial types.

Response time facilitation in response to cross-modal stimuli compared to unimodal auditory ones was observed in both groups. In the simple detection task, RTs to cross-modal stimuli were shorter than RTs to unimodal stimuli by an average of 58ms. However, this cross-modal effect was significantly larger in the auditory modality than in the visual modality. Older adults showed faster RTs to unimodal visual and cross-modal stimuli averaging at 362ms and 352ms, respectively, compared to unimodal auditory RTs, which averaged at 465ms (Table 1a). In the semantic classification task, there was also a significant difference between cross-modal and unimodal auditory trial types: RTs to cross-modal presentations were significant faster than RTs to auditory-only presentations, but not visual-only presentations. Older adults showed a larger difference in performance compared to young adults. They were differentially faster than young adults to respond to cross-modal relative to unimodal auditory presentations by an average of 158ms (Table 1b).

It is possible that older adults exhibit a larger RT gain to cross-modal relative to unimodal auditory presentations because they are slower across all trial types to begin with. In other words, their response latencies may fall further than young adults' during multisensory contexts because the latter group is already significantly faster. Therefore, in order to rule out

the effects of sensorimotor slowing in older compared to young adults, we regressed out age from RTs across both conditions and all three trial types. Standardized residuals were extracted after regressing out age. To capture behavioural facilitation effects in response to cross-modal stimuli, cross-modal RTs were subtracted from unimodal ones for both auditory and visual modalities. This procedure was performed for each condition. After computing the standardized residuals and subtracting cross-modal from unimodal trial types, a one-way ANOVA was used to compare group differences in multisensory facilitation. Mean and standard deviation values of the multisensory facilitation index ($RT_{unimodal}$ - $RT_{cross\text{-}modal}$) are included in Tables 2a and b for the simple detection and the semantic classification conditions, respectively.

3.1.1 Simple detection task

With respect to response latencies, older adults were overall slower than young adults by an average of 107ms. However, significant differences between young and older adults persisted after regressing out age. In response to cross-modal compared to unimodal auditory presentations, older adults were significantly faster than young adults [$F(1, 29) = 7.13$, $p < 0.01$, $\eta^2 = 0.291$] by an average of 38ms. Furthermore, there was a trend towards significant differences between unimodal visual and cross-modal trial types [$F(1,29) = 3.60$, $p = 0.07$, $\eta^2 = 0.113$] in the older group, suggesting that older adults were also faster to respond to cross-modal stimuli compared to unimodal visual ones by an average of 19ms. It is important to note that in the simple detection task, RTs to congruent and incongruent cross-modal pairs were not significantly different in both groups. This suggest that simultaneous presentation of complex sounds and visual objects speeds up response latencies irrespective of the semantic incongruences across sensory modalities during the simple detection task, in which subjects are required to detect any stimulus type as quickly as possible.

3.1.2 Semantic classification task

Similar to the detection condition, older adults were slower than younger adults by an average of 102ms. Additionally, in the semantic classification condition, large differences between young and older groups were observed in the auditory modality. RTs to cross-modal compared to unimodal auditory presentations were significantly shorter in older adults [$F(1, 29) = 21.61$, $p < 0.001$, $\eta^2 = 0.601$] by an average of 156ms. No differences between young and older adults were captured in the visual modality [$F(1, 29) = 0.85$, $p = 0.36$]. Unlike the

behavioural trends in the auditory modality, RTs to cross-modal presentations were not facilitated by concurrent visual presentations in both young and older groups.

Both young and older adults showed reduced differences in performance to cross-modal compared to unimodal visual trial types. In the semantic classification task in particular, the presentation of the visual target facilitated auditory object recognition, but not vice versa. Indeed, previous studies demonstrated an asymmetry between auditory and visual modalities with increased dominance of vision over auditory perception during object recognition tasks similar to those used in the present study (cf. Jaekl & Harris, 2009; Yuval-Greenberg & Deouell, 2009). The additional increase in response latencies in response to complex sounds in the older group may relate to the auditory processing deficits observed in normal aging. Although we controlled for hearing loss and presbycusis, the effects of visual dominance over auditory perception may be more pronounced in older age.

3.2 MEG source activity

To examine age differences in multisensory integration, we used the contrast-driven, non-rotated PLS analysis approach and tested for age-related differences in multisensory responses in the simple detection task. Neural responses to AV+ and AV- trial types did not differ significantly ($p = 0.136$), therefore, we only used the congruent or AV+ trial types in our comparisons of multisensory and unisensory responses.

3.2.1 Simple detection task

Group non-rotated PLS analysis yielded two significant LVs, which reflected multisensory responses or significant differences between cross-modal and unimodal trial types (auditory and visual only) (LV1 = 49.42, $p < 0.05$; LV2 = 70.52, $p < 0.001$). The first LV captured multisensory responses in both young and older adults between 100 and 250ms in the left ventral temporal cortex and the right cuneus (Fig. 2). Conversely, the task effect captured by second LV showed evidence of enhanced multisensory responses in bilateral parietal and medial prefrontal cortices in the older group only. In contrast to young participants, older adults showed increased activity in posterior parietal sources, including the left inferior parietal cortex and the right superior parietal cortex between 150 and 250ms. Enhanced activity in response to cross-modal compared to unimodal trial types was observed in the left medial prefrontal cortex between 150 and 300ms and in the right dorsomedial prefrontal cortex between 250 and 380ms after stimulus onset (Fig. 3).

As the paradigm involved manual responses, the task effects obtained could potentially be confounded at latencies when pre-motor activity may be present (cf. Besle et al., 2004; Teder-Sälejärvi et al., 2002). To address this potential confound, we subdivided trials according to response latencies and examined task-dependent effects in medial prefrontal cortices when RTs did not differ between the two groups. We observed increased MEG source activity in medial prefrontal areas in response to cross-modal compared to unimodal trial types suggesting that the multisensory responses obtained in the older group are not confounded by age differences in RTs. Furthermore, we also examined whether pre-motor sources exhibited significant group differences in multisensory responses. We did not find any significant group by task interactions in pre-motor regions.

3.2.2 Semantic classification task

In the semantic classification task, we observed significant visual dominance effects with RTs to cross-modal and unimodal visual presentations significantly faster than RTs to unimodal auditory presentations. Thus, in this condition, in order to examine age differences in sensory dominance, we contrasted cross-modal to unimodal visual or unimodal auditory trial types separately using mean-centered PLS.

Group mean-centered PLS analysis of cross-modal and unimodal auditory trial types in young and older groups produced two significant LVs (LV1 = 43.79, $p < 0.01$; LV2=40.04, $p < 0.0332$) (Fig. 4a and 4c). The first LV emphasized similarities between the young and older groups, while the second LV captured group differences. In both young and older groups, increased amplitude modulations in response to cross-modal compared to unimodal auditory trial types were observed in the left superior temporal cortex between 100-300ms (Fig. 4b). The second LV captured group differences across the two trial types (Fig. 4c). While young adults showed a lack of task differences, older adults exhibited increased amplitude modulations in response to cross-modal compared to unimodal stimuli in the medial anterior cingulate cortex between 100 and 200ms. Furthermore, older adults showed enhanced activity in the left medial prefrontal cortex in response to cross-modal stimuli between 100 and 300ms, whereas younger adults showed a trend towards increased activity in the medial PFC in response to auditory stimuli compared to cross-modal ones. This trend was not reliable by bootstrap resampling (Fig. 4d).

Mean-centered PLS analysis of cross-modal and unimodal visual trial types produced one significant LV (LV1 =46.71; $p < 0.036$) and captured group similarities in amplitude

modulations (Fig. 5a). Both young and older participants showed increased activity in the right fusiform gyrus in response to cross-modal compared to unimodal visual trial types (Fig. 5b).

3.3 Multisensory responses and Behavioural Performance

To examine whether there was a significant relationship between MEG source activity in posterior parietal or medial prefrontal cortices and cross-modal RT facilitation, we correlated RTs to MEG source activity across all 72 regions. As multisensory facilitation across sensory modalities was only observed in the simple detection task, we examined correlations between RTs and MEG source activity in the simple detection task only. Behaviour PLS analysis revealed a significant relationship between the task differences in RTs and source amplitude modulations in cross-modal, unimodal visual, and unimodal auditory trial types (LV1=141.45, $p < 0.0001$; see Figure 6a). While brain-behaviour correlation patterns did not differ between the three trial types in the younger group, older adults showed robust differences in brain-behaviour correlations in response to cross-modal stimuli compared to unimodal auditory and unimodal visual ones. In the older group, increases in amplitude modulations in the right inferior parietal cortex correlated with faster RTs between 200 and 600ms following cross-modal presentations. Furthermore, enhanced source activity in the right medial PFC between 250 and 450ms also correlated with faster RTs in older compared to younger adults (Fig. 6b). Similar brain-behaviour correlations patterns were observed bilaterally in other posterior parietal and frontal brain regions (Fig. 6c).

3.4 Structural MRI measures

We also investigated how age-related differences in cortical and subcortical volumes related to changes in multisensory responses with age. We used the mean-centered PLS analysis approach to examine the effects of age on white matter and gray matter volumes in cortical and subcortical regions. Significant differences between young and older adults were observed (LV1=3.86, $p < 0.0001$; Fig. 7a) such that, in comparison to young adults, older adults showed significantly reduced gray matter volumes across frontal, parietal, temporal and occipital lobes. In terms of age differences in white matter volumes, significant reductions were only captured in white matter pathways in the frontal lobes. Reductions in gray matter volumes were also observed in subcortical structures, including the striatum, thalamus, hippocampus, and cingulate cortices (Fig. 7b).

3.4 .1 Structural MRI measures and Behavioural Performance

To examine the relationship between the structural changes that occur with age and multisensory facilitation in simple detection and classification conditions, we correlated the RT differences in both auditory and visual modalities with gray and white matter volumes in cortical and sub-cortical regions in young and older adults. Corrections for multiple comparisons were performed using the Bonferroni test.

The correlation patterns, which survived multiple comparison correction, were observed in several cortical regions (incl. the frontal, temporal, occipital and parietal lobes). Faster RTs to cross-modal stimuli relative to unimodal auditory ones correlated with reductions in gray matter volumes ($r = -0.47$, $p < 0.01$; $r = -0.63$, $p < 0.01$) in simple detection and semantic classification conditions, respectively. In the visual modality, shorter RTs to cross-modal relative to unimodal visual stimuli correlated with reductions in gray matter volumes ($r = -0.42$, $p < 0.05$) in the simple detection condition. In the semantic classification condition, a larger RT gain to cross-modal compared to unimodal auditory trial types correlated with decreases in gray matter volumes in the (i) thalamus ($r = -0.59$, $p < 0.01$), (ii) bilateral medial temporal lobes ($r = -0.61$, $p < 0.01$) and (iii) bilateral anterior and posterior cingulate cortices ($r = -0.62$, $p < 0.01$) in older compared to younger adults.

As RT facilitation in response to cross-modal stimuli was significantly correlated with multisensory responses in parietal and medial prefrontal regions, as well as to age-related reductions in gray and white matter volumes, we also assessed the relationship between all three dependent variables, namely cross-modal RT facilitation, volumetric measures, and MEG source activity in the two age groups, using SEM.

3.5 SEM Results

Pair-wise comparisons between the two groups within each condition indicated a significant difference in the model fit in the two tasks in response to cross-modal stimuli compared to both auditory [$\chi^2_{DIFF}(15) = 79.25$ ($p < 0.0001$)] and visual stimuli [$\chi^2_{DIFF}(15) = 64.76$ ($p < 0.001$)]. Refer to Table 3 for the individual path coefficients and to Figure 8 for a schematic of the directional effects between RT facilitation and the structural and functional measures (i.e., task-related MEG source activity) in the two age groups across both experimental conditions.

In the older group, cortical and sub-cortical gray matter volume decline predicted larger multisensory responses in parietal and medial prefrontal sources. Furthermore, reductions in gray matter volumes with age also predicted faster RTs to cross-modal compared to unimodal auditory stimuli. No differences in the path coefficients were observed in the younger group (Fig. 8).

In the visual modality, however, both young and older adults failed to show significant interactions between the three variables of interest. In the younger group, however, there was a trend towards a significant relationship between volumetric measures and MEG source activity indicating that larger gray and white matter volumes predicted reduced differences in amplitude modulations between cross-modal and unimodal visual trial types (cf., Table 3).

4. Discussion

Young and older adults exhibited enhanced cross-modal facilitation relative to the auditory modality. Behaviourally, both groups were faster to classify complex sounds when they were paired with visual stimuli. Older adults, however, showed a more pronounced gain in performance during cross-modal trials compared to unimodal auditory ones. Using MEG recordings and contrasting cross-modal presentations to both unimodal counterparts, we demonstrated age-related differences in multisensory responses in the simple detection task. Larger multisensory responses in posterior parietal and medial prefrontal cortices also predicted faster performance in response to cross-modal events in older adults. Reductions in gray matter volumes across cortical and sub-cortical regions were associated with enhanced multisensory responses in posterior parietal and medial prefrontal regions in older adults, and jointly predicted faster performance to cross-modal stimuli. Similarly to the behavioural results, these effects were more pronounced in the auditory modality suggesting that visual dominance effects are exacerbated with age.

4. 1 Visual Dominance

The behavioural findings in the simple detection task suggest that older adults show a trend towards enhanced RT facilitation to cross-modal compared to unimodal stimuli. However, this cross-modal RT facilitation effect was larger in the auditory modality suggesting that older participants were faster to respond to complex sounds when they were accompanied by visual stimuli than vice versa. A more pronounced influence of concurrent visual stimulation on auditory perception has been demonstrated in previous behavioural studies (cf., Chen & Spence, 2010; Yuval-Greenberg & Deouell, 2009). Faster RTs to tones following concurrent AV presentation may occur because the visual modality provides more reliable and unambiguous information for object recognition. This interpretation is also consistent with the documented situations of visual dominance over auditory processing as in the Colavita effect, which suggests that the detection of auditory targets is significantly reduced when tones are presented alongside visual stimuli in multisensory conditions. Conversely, in the visual modality, the detectability of visual targets was significantly assisted by concurrent auditory presentations (Colavita, 1974).

An even larger effect of sensory modality was noted in the semantic classification condition, in which participants were required to judge whether the unimodal and cross-modal stimuli belonged to animate or inanimate categories. Auditory object categorization was

enhanced following complementary presentations of visual stimuli; however, no significant differences between unimodal visual and cross-modal stimulus presentations were detected in this task suggesting that the visual sensory modality was more dominant than the auditory modality, and, therefore, performance in response to visual stimuli did not benefit from concurrent presentations of complex sounds. This visual sensory dominance effect may be more pronounced in this condition due to the nature of the semantic classification task. When required to identify the class to which complex sounds belonged, participants need to temporally integrate the sounds to accumulate enough evidence about their respective category. In contrast to the auditory modality, the detection of the category of the visual stimulus is not graded, but virtually instantaneous. In order to maintain ecological validity, we presented the stimuli without manipulating the visual images to match the temporal integration of the complex sounds, thus maintaining the sensory modality bias.

Multisensory relative to auditory object categorization was significantly better with age after ruling out potential hearing impairments such as presbycusis. Older adults may show a more pronounced visual dominance effect because they have more difficulties when performing listening tasks that require temporal processing compared to young adults (cf., Pichora-Fuller et al., 1995; Schneider et al., 1998). This effect, however, is reduced when older participants are instructed to actively attend to the temporal structure of the auditory presentations. Alain and colleagues showed that while automatic central auditory processing may be reduced in older adults (Alain, McDonald, Ostroff, & Schneider, 2004; for review: Alain, Dyson, & Snyder, 2006), older adults are as sensitive as young adults to near-threshold deviance gaps during active listening conditions.

4. 2 Multisensory Integration

Multisensory responses were captured in both groups in sensory specific brain regions, namely the superior and ventral temporal cortices and the cuneus, between 100 and 300ms. Such amplitude modulations extending 100ms after the onset of cross-modal stimuli were observed in both the simple detection and the semantic classification task. In a recent study, we examined multisensory processes in response to both cross-modal congruent and incongruent stimuli and found evidence of enhanced multisensory integration in sensory-specific and posterior parietal sources within 100ms after stimulus onset (Diaconescu et al., 2011).

Previous ERP studies employing simple detection tasks, such as those used in the present study, also demonstrated significant amplitude deflections in centro-temporal channels over

auditory cortices and in posterior-occipital channels over visual cortices between 100 and 200 ms (cf., Giard & Peronnet, 1999; Teder-Salejarvi et al., 2005; Teder-Salejarvi, McDonald, Di Russo, & Hillyard, 2002). While the behavioural results in the semantic classification condition indicate dominance of vision over audition, previous ERP findings suggest that auditory animacy judgments begin within 100ms after stimulus onset (Murray, Camen, Gonzalez Andino, Bovet, & Clarke, 2006), and such amplitude modulations are also within the timeframe attributed to animacy judgments of visual objects (Thorpe, Fize, & Marlot, 1996).

In a recent study examining the effects of multisensory integration in young and older adults, Stephen and colleagues (2010) used a similar approach to investigate age-related multisensory facilitation effects. The group found evidence of multisensory facilitation in response to cross-modal stimuli relative to unimodal ones and reported a behavioural trend towards larger RT gains in response to multisensory stimuli in the older group. Similarly to the present study, the authors used MEG recordings and tested 8 young and older participants in an AV spatial congruency paradigm. A picture of a soccer ball was presented in two locations along the vertical meridian: *near* versus *far* relative to the observer. The soccer ball was presented along with two distinct types of tones: a faint tone associated with the farther visual stimulus, and a loud tone associated with the closer visual stimulus. Similar to the present study, visual dominance effects were detected and reduced multisensory facilitation effects were observed in conditions in which visual stimuli were paired with concurrent auditory presentations. The group performed source localization using multi-dipole spatiotemporal modeling and detected larger amplitude modulations in the superior temporal cortex, the insula, parietal and cingulate regions to cross-modal presentations compared to unimodal ones in both young and older groups. However, the authors did not detect any evidence of age-related differences in multisensory responses (Stephen, Knoefel, Adair, Hart, & Aine, 2010).

4. 2 Age Differences in Multisensory Integration and Visual Dominance

In addition to age differences in behavioral facilitation, we also found age differences in the spatiotemporal dynamics in response to cross-modal stimuli compared to unimodal auditory and unimodal visual ones. Comparisons across young and older groups revealed significant differences in multisensory responses in the simple detection task. Age differences in multisensory integration were not captured in basic sensory areas, but in multisensory posterior parietal and frontal sources in the simple detection task. Source activity in posterior parietal and

medial prefrontal regions was larger in response to cross-modal stimuli relative to both unimodal visual and unimodal auditory ones.

We predicted that faster performance to cross-modal compared to unimodal presentations reflects a change in functional network organization beyond sensory-specific channels encompassing multisensory prefrontal areas. Indeed, older adults activated a distinct network of brain regions in response to cross-modal presentations. In the simple detection condition, age differences in multisensory responses were captured in parietal and medial prefrontal MEG sources, including the inferior and superior parietal cortex between 150 and 300ms and the bilateral medial prefrontal cortices between 150-250 and 200-300ms. Furthermore, increased source activity in posterior parietal and prefrontal regions also predicted faster performance to cross-modal compared to both unimodal auditory and visual stimuli in the simple detection task.

In the semantic classification task, visual dominance effects were observed 100ms after stimulus onset with participants in the elderly group showing enhanced activations in cingulate and medial prefrontal regions in response to cross-modal stimuli compared to unimodal auditory ones; however, no differences were detected between cross-modal and unimodal visual stimuli. Similar to the behavioural results in this task, cross-modal facilitation was larger relative to the auditory modality compared to the visual modality suggesting that visual dominance effects are exacerbated with age.

Participants who showed enhanced multisensory responses in parietal and prefrontal sources also exhibited more pronounced reductions in cortical and sub-cortical gray matter volumes. In line with previous research on gray matter volume changes with age (cf., Gunning-Dixon & Raz, 2000; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003), in spite of equivalent performance on neuropsychological tests, older participants exhibited reduced gray matter volumes in both cortical and sub-cortical regions relative to younger adults. Reductions in cortical and sub-cortical gray matter volumes predicted larger multisensory responses in parietal and medial prefrontal sources, which in turn, predicted enhanced cross-modal RT facilitation. This effect, however, was more pronounced in the auditory modality suggesting a larger visual dominance effect in older adults compared to younger. The relationship between the performance trends and the source activity patterns was primarily mediated by the age-related reductions in gray matter volume. In other words, older adults who were faster to respond to cross-modal stimuli compared to unimodal auditory ones also showed reduced

cortical and sub-cortical gray matter volumes and recruited a distinct network of posterior parietal and medial prefrontal sources in response to cross-modal compared to unimodal trial types.

By examining the neural mechanisms that support multisensory integration and visual sensory dominance, we propose that age-related differences in responses to cross-modal stimuli are a consequence of functional reorganization, which, in turn, arises from structural declines with age. Previous neuroimaging studies showed that additional recruitment of parietal and prefrontal resources was functionally adaptive in older adults, and led to improvements in performance. Bilateral recruitment of parietal and prefrontal cortices was primarily observed in older participants who performed better or as well as young adults on a variety of perceptual and cognitive tasks (Cabeza et al., 2002; Grady et al., 1994; Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Cappell, 2008; Thorpe et al., 1996). Furthermore, inhibition of prefrontal activity during recognition memory tasks significantly impaired older adults' performance (Rossi et al., 2004), whereas stimulation of prefrontal activity in under-performing older adults improved memory scores (Sole-Padulles et al., 2006) suggesting an adaptive role of prefrontal recruitment in healthy aging.

Our results indicate that both multisensory integration and visual dominance are more pronounced with age. Enhanced multisensory responses in posterior parietal and medial prefrontal regions may serve a compensatory function as they predict cross-modal facilitation in older adults. Such compensatory effects may arise as a result of gray matter volume changes and reductions in temporal processing during auditory perception that accompany healthy aging. Future research is required to further examine whether multisensory integration or visual dominance can confer a benefit in older adults who exhibit significant motor or cognitive impairments. An extension of the present work is to determine whether cognitive performance or motor vigilance can be improved with the addition of cross-modal cues.

5. References:

Alain, C., Dyson, B. J., & Snyder, J. S. (2006). Aging and the perceptual organization of sounds: A change of scene? *Handbook of Models for Human Aging* (pp. 759-769).

Alain, C., McDonald, K. L., Ostroff, J. M., & Schneider, B. (2004). Aging: A Switch From Automatic to Controlled Processing of Sounds? *Psychology and Aging*, *19*(1), 125–133

Alain, C., Ogawa, K. H., & Woods, D. L. (1996). Aging and the segregation of auditory stimulus sequences. *J Gerontol B Psychol Sci Soc Sci, 51*(2), P91-93.

Baumann, O., & Greenlee, M. W. (2007). Neural correlates of coherent audiovisual motion perception. *Cereb Cortex, 17*(6), 1433-1443.

Besle, J., Fort, A., Delpuech, C., & Giard, M.-H. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *European Journal of Neuroscience*, *20*(8).

Bezgin, G., Wanke, E., Krumnack, A., & Kotter, R. (2008). Deducing logical relationships between spatially registered cortical parcellations under conditions of uncertainty.. *Neural networks, 21*(8), 1132-1145.

Bishop, C. W., & Miller, L. M. (2008). A Multisensory Cortical Network for Understanding Speech in Noise. *J Cogn Neurosci*.

Brookes, M. J., Stevenson, C. M., Barnes, G. R., Hillebrand, A., Simpson, M. I., Francis, S. T., & Morris, P. G. (2007). Beamformer reconstruction of correlated sources using a modified source model. *Neuroimage, 34*(4), 1454-1465.

Bunce, D. J., Warr, P. B., & Cochrane, T. (1993). Blocks in choice responding as a function of age and physical fitness. *Psychol Aging, 8*(1), 26-33.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage, 17*(3), 1394-1402.

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage, 14*(2), 427-438.

Campbell, K. L., Al-Aidroos, N., Fatt, R., Pratt, J., & Hasher, L. (2010). The effects of multisensory targets on saccadic trajectory deviations: eliminating age differences. *Exp Brain Res, 201*(3), 385-392.

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: a unique age effect. *Psychol Sci, 21*(3), 399-405.

Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2010). Auditory–Visual Multisensory Interactions in Humans: Timing, Topography, Directionality, and Sources. *The Journal of Neuroscience*, *30*(38), 12572–12580.

Carp, J., Park, J., Polk, T. A., & Park, D. C. (2011). Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *Neuroimage, 56*(2), 736-743.

Cerella, J. (1985). Age-related decline in extrafoveal letter perception. *J Gerontol, 40*(6), 727-736.

Chen, Y. C., & Spence, C. (2010). When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures. *Cognition, 114*(3), 389-404.

Cheyne, D., Bakhtazad, L., & Gaetz, W. (2006). Spatiotemporal mapping of cortical activity accompanying voluntary movements using an event-related beamforming approach. *Hum Brain Mapp, 27*(3), 213-229.

Cheyne, D., Bostan, A. C., Gaetz, W., & Pang, E. W. (2007). Event-related beamforming: a robust method for presurgical functional mapping using MEG. *Clin Neurophysiol, 118*(8), 1691-1704.

Colavita, F. B. (1974). Human sensory dominance. *Percept Psychophys, 16*(2), 409-412.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res, 29*(3), 162-173.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends Cogn Sci, 10*(3), 131-138.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179-194.

Diaconescu, A. O., Alain, C., & McIntosh, A. R. (2011). The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *J Neurophysiol, 106*(6), 2896-2909.

Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia, 46*(10), 2556-2562.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science, 1*, 54-77.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res, 12*, 189-198.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci, 11*(5), 473-490.

Grady, C. L., & Craik, F. I. (2000). Changes in memory processing with age. *Curr Opin Neurobiol, 10*(2), 224-231.

Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., . . . Haxby, J. V. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *J Neurosci, 14*(3 Pt 2), 1450-1462.

Grefkes, C., Weiss, P. H., Zilles, K., & Fink, G. R. (2002). Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron, 35*(1), 173-184.

Guerreiro, M. J. S., Murphy, D. R., & Van Gerven, P. W. M. (2010). The role of sensory modality in age-related distraction: A critical review and a renewed view. *Psychological Bulletin*, *136*(6), 975–1022.

Gunning-Dixon, F. M., & Raz, N. (2000). The cognitive correlates of white matter abnormalities in normal aging: a quantitative review. *Neuropsychology, 14*(2), 224-232.

Hartman, M., & Hasher, L. (1991). Aging and suppression: memory for previously relevant information. *Psychol Aging, 6*(4), 587-594.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a Short Orientation-Memory-Concentration Test of Cognitive Impairment. *American Journal of Psychiatry, 140*(6), 734-739.

Kotter, R., & Wanke, E. (2005). Mapping brains without coordinates. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 751-766.

Krishnan, A., Williams, L. J., McIntosh, A. R., & Abdi, H. (2011). Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage, 56*(2), 455-475.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol Aging, 27*(8), 1155-1163.

Lobaugh, N. J., West, R., & McIntosh, A. R. (2001). Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. *Psychophysiology, 38*(3), 517-530.

Lustig, C., Hasher, L., & Tonev, S. T. (2006). Distraction as a determinant of processing speed. *Psychon Bull Rev, 13*(4), 619-625.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage, 21*(2), 725-732.

McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. (1996). Spatial pattern analysis of functional brain images using Partial Least Squares. *Neuroimage, 3*, 143-157.

McIntosh, A. R., & Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping, 2*(1-2), 2-22.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res, 14*(1), 115-128.

Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., . . . Foxe, J. J. (2006). Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *J Neurophysiol, 96*(2), 721-729.

Moran, R. J., Molholm, S., Reilly, R. B., & Foxe, J. J. (2008). Changes in effective connectivity of human superior parietal lobule under multisensory and unisensory stimulation. *Eur J Neurosci, 27*(9), 2303-2312.

Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *J Neurosci, 26*(4), 1293-1302.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp, 15*(1), 1-25.

Okada, Y., Jung, J., & Kobayashi, T. (2007). An automatic identification and removal method for eye-blink artifacts in event-related magnetoencephalographic measurements. *Physiol Meas, 28*(12), 1523-1532.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. [Research Support, U.S. Gov't, P.H.S.]. *Psychol Aging, 17*(2), 299-320.

Park, D. C., Polk, T. A., Park, R., Minear, M., Savage, A., & Smith, M. R. (2004). Aging reduces neural specialization in ventral visual cortex. *Proc Natl Acad Sci U S A, 101*(35), 13091-13095.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu Rev Psychol, 60*, 173-196.

Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport, 18*(10), 1077-1081.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J Acoust Soc Am, 97*(1), 593-608.

Poliakoff, E., Coward, R. S., Lowe, C., & O'Boyle, D. J. (2007). The effect of age on inhibition of return is independent of non-ocular response inhibition. *Neuropsychologia, 45*(2), 387-396.

Poliakoff, E., Shore, D. I., Lowe, C., & Spence, C. (2006). Visuotactile temporal order judgments in ageing.. *Neurosci Lett, 396*(3), 207-211.

Protzner, A. B., & McIntosh, A. R. (2006). Testing effective connectivity changes with structural equation modeling: what does a bad model tell us? *Hum Brain Mapp, 27*(12), 935-947.

Quraan, M. A., & Cheyne, D. (2010). Reconstruction of correlated brain activity with adaptive spatial filters in MEG. *NeuroImage*, *49*(3), 2387–2400.

Raij, T., Ahveninen, J., Lin, F.-H., Witzel, T., Jääskeläinen, I. P., Letham, B., Israeli, E., et al. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, *31*(10).

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J Neurosci, 23*(8), 3295-3301.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science, 17*(3), 177-182.

Robinson, S., & Vrba, J. (Eds.). (1998). *Functional neuroimaging by synthetic aperture magnetometry*. Sendai: Tohoku Univ. Press.

Rossi, S., Miniussi, C., Pasqualetti, P., Babiloni, C., Rossini, P. M., & Cappa, S. F. (2004). Age-related functional changes of prefrontal cortex in long-term memory: a repetitive transcranial magnetic stimulation study. *J Neurosci, 24*(36), 7939-7944.

Salthouse, T. A. (2007). What and When of Cognitive Aging. *Current Directions in Psychological Science, 13*(4), 140-144.

Schneider, B., Speranza, F., & Pichora-Fuller, M. K. (1998). Age-related changes in temporal resolution: envelope and intensity effects. *Can J Exp Psychol, 52*(4), 184-191.

Scialfa, C. T., Esau, S. P., & Joffe, K. M. (1998). Age, target-distractor similarity, and visual search. *Exp Aging Res, 24*(4), 337-358.

Sekihara, K., Nagarajan, S. S., Poeppel, D., Marantz, A., & Miyashita, Y. (2001). Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. *IEEE Trans Biomed Eng, 48*(7), 760-771.

Setti, A., Finnigan, S., Sobolewski, R., McLaren, L., Robertson, I. H., Reilly, R. B., . . . Newell, F. N. (2011). Audiovisual temporal discrimination is less efficient with aging: an event-related potential study. *Neuroreport, 22*(11), 554-558.

Shipley, W. C. (1991). *Shipley Institute of Living Scale*. Los Angeles: Western Psychological Services.

Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision Res, 33*(18), 2589-2609.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *J Exp Psychol Hum Percept Perform, 22*(2), 461-479.

Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., & Aine, C. J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neurosci Lett, 484*(1), 76-80.

Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perrault, T. J., Ramachandran, R., et al. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *European Journal of Neuroscience*, *31*(10), 1713–1720.

Stoltzfus, E. R., Hasher, L., Zacks, R. T., Ulivi, M. S., & Goldstein, D. (1993). Investigations of inhibition and interference in younger and older adults. *J Gerontol, 48*(4), P179-188.

Sole-Padulles, C., Bartres-Faz, D., Junque, C., Clemente, I. C., Molinuevo, J. L., Bargallo, N., et al. (2006). Repetitive transcranial magnetic stimulation effects on brain function and cognition among elders with memory dysfunction. A randomized sham-controlled study. *Cereb Cortex, 16*(10), 1487-1493.

Teder-Sälejärvi, W. A., Russo, F. D., McDonald, J. J., & Hillyard, S. A. (2005). Effects of Spatial Congruity on Audio-Visual Multimodal Integration. *Journal of Cognitive*

Neuroscience, *17*(9), 1396–1409.
Teder-Sälejärvi, W. A., McDonald, J. ., Di Russo, F., & Hillyard, S. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, *14*(1), 106–114.
Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520-522.
West, R., & Alain, C. (1999). Event-related neural activity associated with the Stroop task. *Brain Res Cogn Brain Res, 8*(2), 157-164.
West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology, 37*(2), 179-189.
Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog's meow: asymmetrical interaction in cross-modal object recognition. *Exp Brain Res, 193*(4), 603-614.
Zeef, E. J., Sonke, C. J., Kok, A., Buiten, M. M., & Kenemans, J. L. (1996). Perceptual factors affecting age-related differences in focused attention: performance and psychophysiological analyses. *Psychophysiology, 33*(5), 555-565.

6. Figure Captions:

Figure 1: Experimental design :Three possible stimulus combinations were used in this study: (1) unimodal visual (V; example, picture of a violin), (2) unimodal auditory (A; example, sound of a "roar"), and (3) cross-modal congruent or simultaneous auditory and visual stimuli that were matched semantically (AV+; example, picture of a bird matched with a corresponding "chirp" sound or AV-; example, picture of a lion paired with the sound of a police siren). Each stimulus or stimulus pair was presented for 400ms; for the auditory stimulus, the 400ms interval also included a 5ms fall and rise time. The time interval between the end of the stimulus presentation and the beginning of the next trial was either 2, 3, and 4s (equiprobable).

Figure 2: PLS results reflecting the multisensory response common across the two groups. The source waveforms were derived using the Talairach coordinates displayed on the left. The area number pertains to the brain region number listed in Supplementary Table 2. The bootstrap ratios (BSRs) below the source waveforms reflect the positive expression of the task contrast [1 -1 -1; 1 -1 -1] - the multisensory response across the two age groups.

Figure 3: PLS results reflecting differences between the two groups in cross-modal compared to both unimodal auditory and unimodal visual trial types. The source waveforms were derived using the Talairach coordinates displayed on the left. The area number pertains to the brain region number listed in Supplementary Table 2. The bootstrap ratios (BSRs) below the source waveforms reflect the positive expression of the task contrast [-1 1 1; 1 -1 -1]- the multisensory response observed in the older group only.

Figure 4: Group mean-centered PLS results in the semantic classification task across cross-modal and unimodal auditory trial types (LV1 and LV2). The task saliences reflect group similarities (a) and group differences (c) in cross-modal and unimodal auditory trial types. The source waveforms were derived using the Talairach coordinates displayed on the left. The area number pertains to the brain region number listed in Supplementary Table 2 (b and d). The bootstrap ratios (BSRs) below the source waveforms reflect the positive expression of the given task saliences (a and c) or increased activity to cross-modal compared to unimodal auditory trial types across the two age groups (a). In (c), they reflect task by group interactions or increased activity to cross-modal compared to unimodal auditory trial types the older group, and increased activity to unimodal auditory compared to cross-modal trial types in the younger group.

Figure 5:
Group mean-centered PLS results in the semantic classification task across cross-modal and unimodal visual trial types (LV1). The task salience reflects group similarities across cross-modal and unimodal visual trial types (a). The source waveforms were derived using the Talairach coordinates displayed on the left. The area number pertains to the

brain region number listed in Supplementary Table 2 (b). The bootstrap ratios (BSRs) below the source waveforms reflect the positive expression of the given task salience (a) or increased activity to cross-modal compared to unimodal visual trial types across the two age groups.

Figure 6: Brain-behaviour correlations were compared across cross-modal and unimodal trial types in the simple detection condition. The task salience in (a) reflects differences in RT and source activity correlations across the cross-modal, unimodal auditory, and unimodal visual trial types. The source saliences in (b) reflect the corresponding brain behaviour correlation pattern in two sources, namely the right inferior parietal cortex and the right medial prefrontal cortex, and across the entire epoch. The source waveforms were computed based on the Talairach coordinates displayed on the left. The area number pertains to the brain region number listed in Supplementary Table 2 and it is overlaid on a template brain. The BSRs below the source waveforms reflect the positive expression of the brain-behaviour correlation pattern in (a) – i.e., negative correlations between RTs and MEG source activity in response to cross-modal stimuli in the older compared to the younger group. Source saliences are also displayed across all 72 brain regions across the time points in which the brain-behaviour correlation pattern is expressed most reliably in (c). The positive BSRs reflect the positive expression of the given task salience in (a).

Figure 7: The task salience in (a) reflects group differences in white and gray matter volumes and cortical thickness measures. BSR values reflect which brain structures reliably express the given task salience (b). Therefore, positive BSRs indicate reduced volumes in older compared to younger adults. This was observed across cortical gray matter, white matter pathways in the frontal lobes, striatum and thalamus structures, and medial temporal and cingulate cortices.

Figure 8: Hypothesis-driven model: Arrows reflect relationships between variables; black arrows reflect significant standardized coefficients and gray arrows reflect non-significant coefficients. The direction of the arrow reflects the hypothesized direction of the relationships. In the older group (dashed lines), increased multisensory responses in parietal and medial prefrontal sources was related to reductions in cortical/sub-cortical gray matter volumes, which in turn, predicted faster RTs to cross-modal compared to unimodal auditory stimuli. Young adults (gray lines) did not show any significant interactions.